FIG. 1

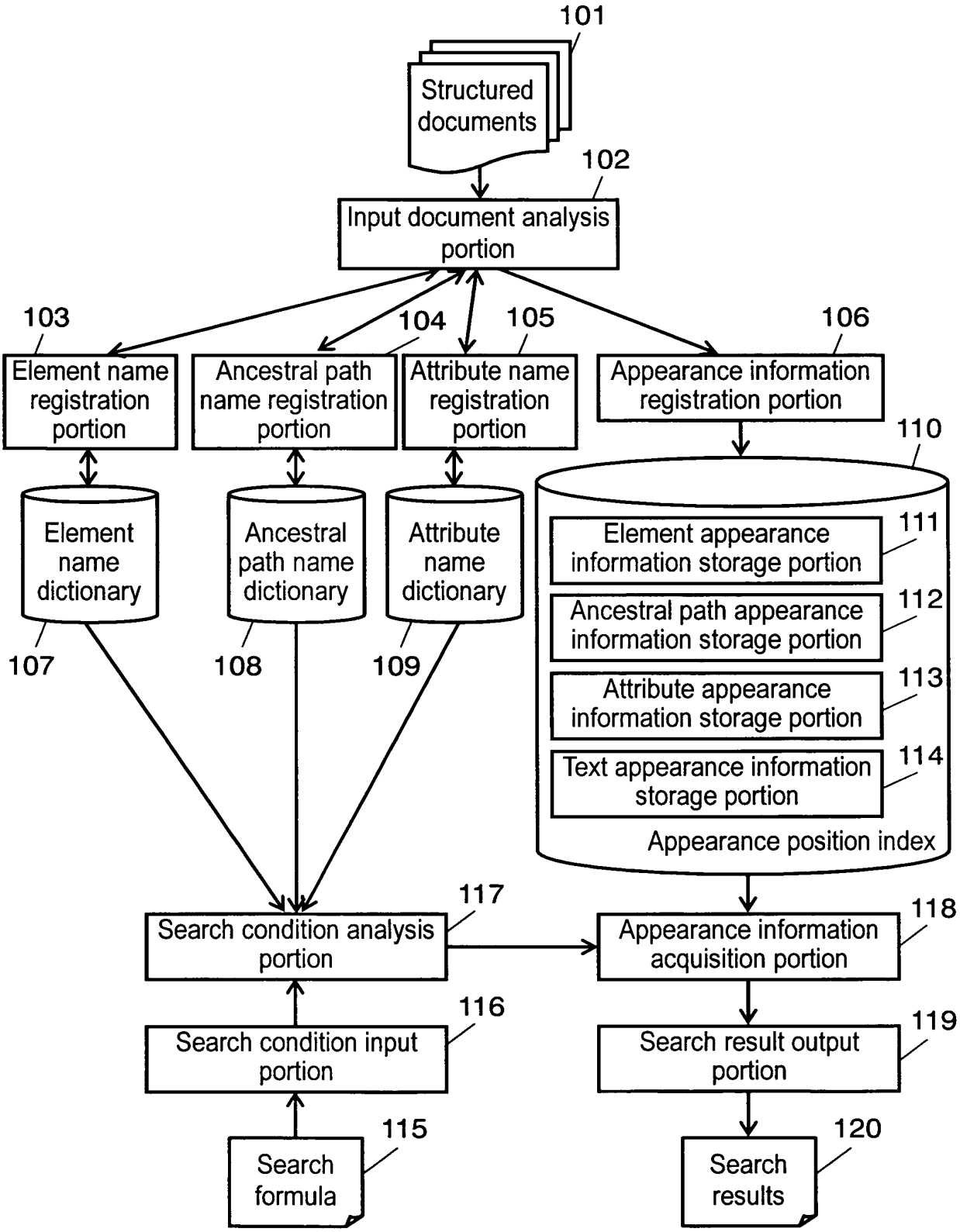
101
Structured documents
102
Input document analysis portion
103
Element name registration portion
104
Ancestral path name registration portion
105
Attribute name registration portion
106
Appearance information registration portion
110
Element name dictionary
Ancestral path name dictionary
Attribute name dictionary
107
108
109
Element appearance information storage portion
111
Ancestral path appearance information storage portion
112
Attribute appearance information storage portion
113
Text appearance information storage portion
114
Appearance position index
117
Search condition analysis portion
118
Appearance information acquisition portion
116
Search condition input portion
119
Search result output portion
115
Search formula
120
Search results

FIG. 2

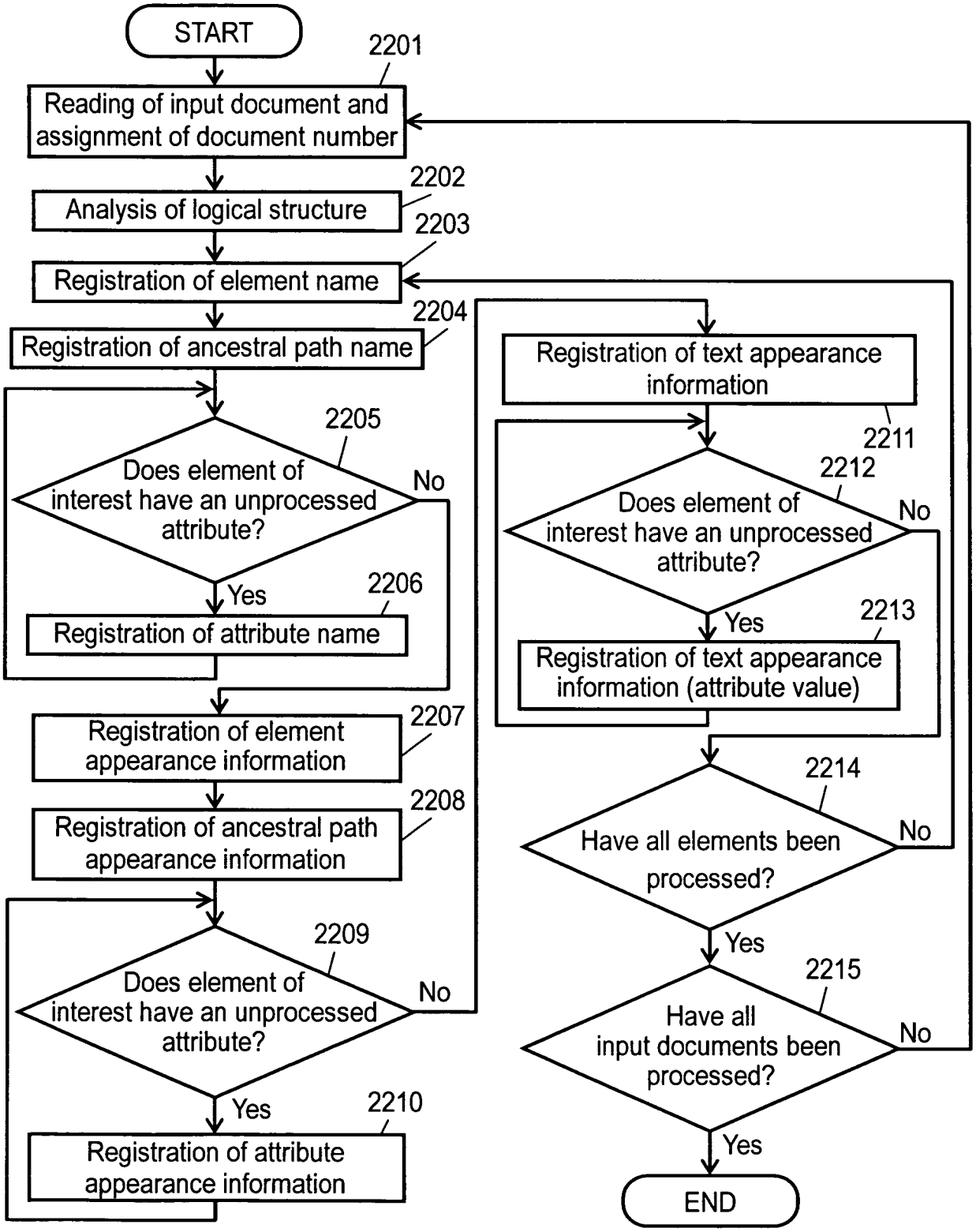
START
2201
Reading of input document and assignment of document number
2202
Analysis of logical structure
2203
Registration of element name
2204
Registration of ancestral path name
2205
Does element of interest have an unprocessed attribute?
No
Yes
2206
Registration of attribute name
2207
Registration of element appearance information
2208
Registration of ancestral path appearance information
2209
Does element of interest have an unprocessed attribute?
No
Yes
2210
Registration of attribute appearance information
Registration of text appearance information
2211
2212
Does element of interest have an unprocessed attribute?
No
Yes
2213
Registration of text appearance information (attribute value)
2214
Have all elements been processed?
No
Yes
2215
Have all input documents been processed?
No
Yes
END

FIG. 3

101a

```
<book>
  <title> document search </title>
  <chapter keyword = "history">
    <title> history of document search </title>
    <section> keyword search is ... </section>
    <section> then, the full text search is ... </section>
  </chapter>
  <chapter keyword = "index">
    <title> search method </title>
    <section> using extraction by longest match principle < /section>
    <section> n-gram search method is ... < /section>
    <section update> "200406" Newly, a maximum word search method is
                                                    </section> </chapter>
  </chapter>
</book>
```

FIG. 4

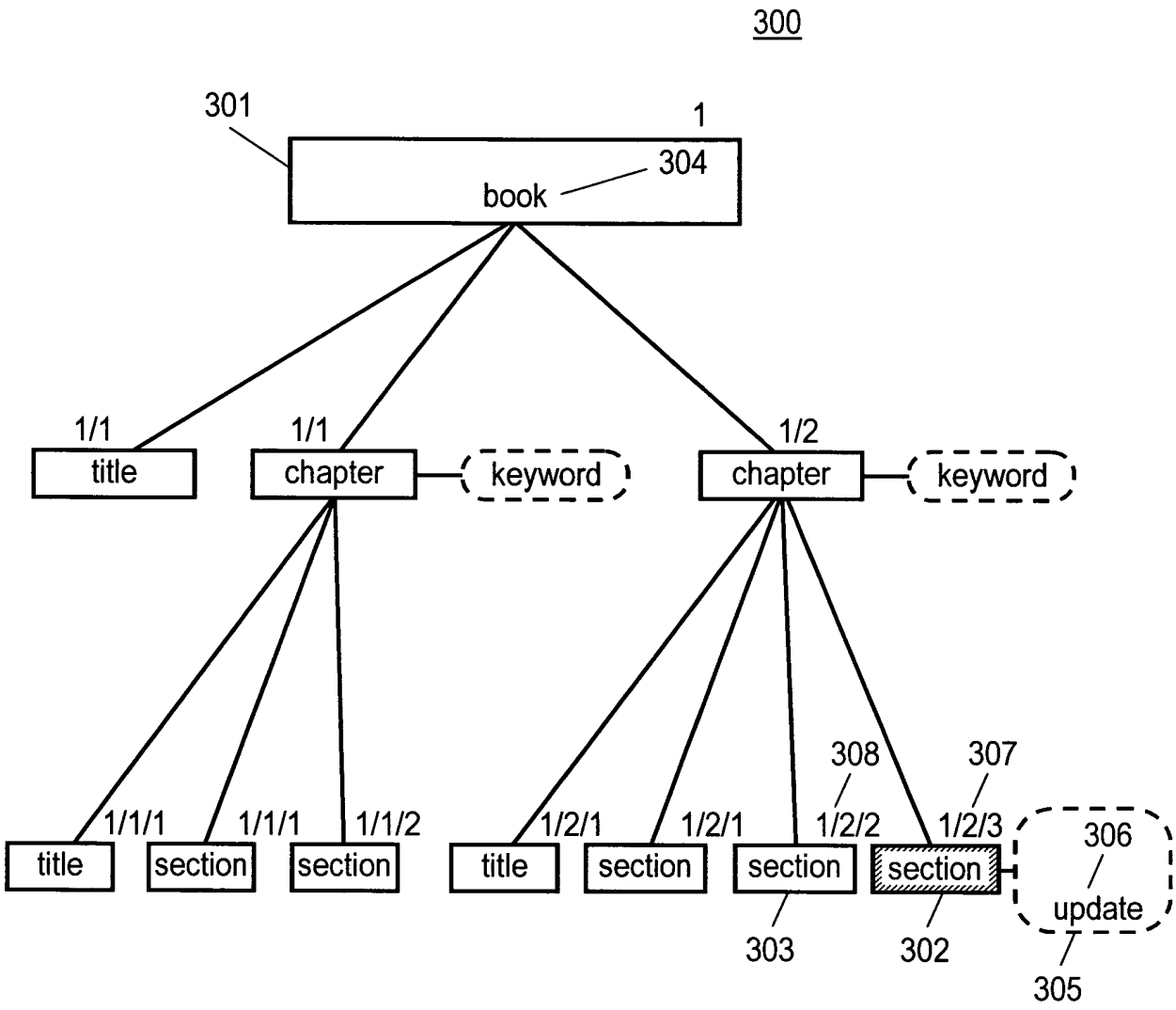
300
301
1
304
book
1/1
title
1/1
chapter
keyword
1/2
chapter
keyword
1/1/1
title
1/1/1
section
1/1/2
section
1/2/1
title
1/2/1
section
308
1/2/2
section
307
1/2/3
section
306
update
303
302
305

FIG. 5

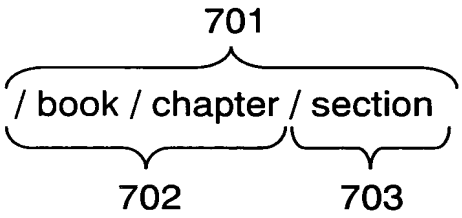

FIG. 6

407

Element name ID	Element name
1	book
2	title
3	chapter
4	section

FIG. 7

408

Ancestral path name ID	Ancestral path name
1	/
2	/ book
3	/ book / chapter

FIG. 8

409

Attribute name ID	Attribute name
1	keyword
2	update

FIG. 9

411

410

Character	docu-	ment	sear-	ch	docu-	ment	sear-	ch	of	his-	tory
Character position	0	1	2	3	4	5	6	7	8	9	10

412

ke	y	w	or	d	sear-	ch	is	,	...
11	12	13	14	15	16	17	18	19	

FIG. 10A

501

502 Element name ID	503 Document No.	504 Character position	505 Number of characters	506 Ancestral path name ID	507 Order of branches
4	1	115	40	3	1/2/3

FIG. 10B

302

322

<section update "200406" >Newly, a maximum word search method is ...</section>

321

FIG. 11

511

506 Ancestral path name ID	503 Document No.	504 Character position	505 Number of characters	502 Element name ID	507 Order of branches
3	1	115	40	4	1/2/3

FIG. 12A

521

522 Attribute name ID	503 Document No.	504 Character position	505 Number of characters	506 Ancestral path name ID	502 Element name ID	507 Order of branches
2	1	115	6	3	4	1/2/3

FIG. 12B

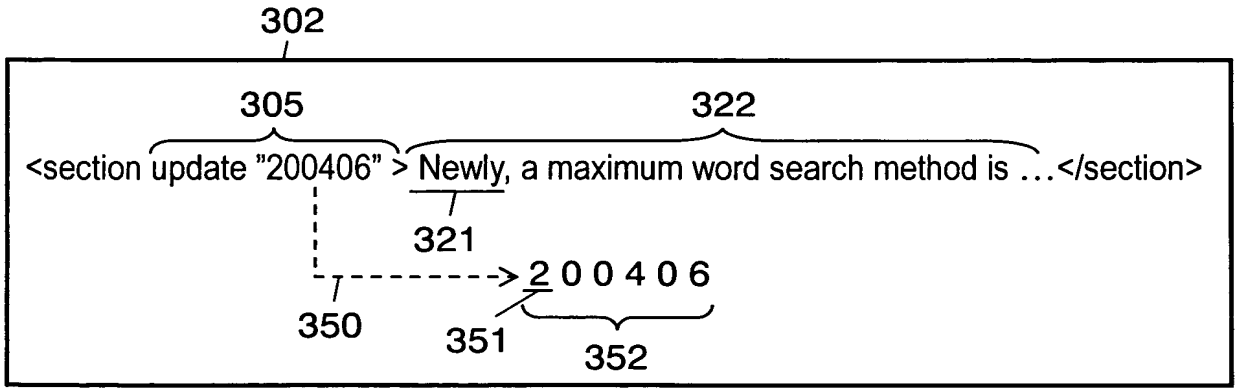

FIG. 13

531

532 Partial character string	503 Document No.	504 Character position	506 Ancestral path name ID	502 Element name ID	522 Attribute name ID	507 Order of branches	
new-	1	115	3	4	0	1/2/3	
ly	1	116	3	4	0	1/2/3	1201
max-	1	117	3	4	0	1/2/3	
maximum	1	118	3	4	0	1/2/3	
-mum word	1	119	3	4	0	1/2/3	
word	1	120	3	4	0	1/2/3	
“20”	1	115	3	4	2	1/2/3	1202
“00”	1	116	3	4	2	1/2/3	
“04”	1	117	3	4	2	1/2/3	
“40”	1	118	3	4	2	1/2/3	
“06”	1	119	3	4	2	1/2/3	

FIG. 14

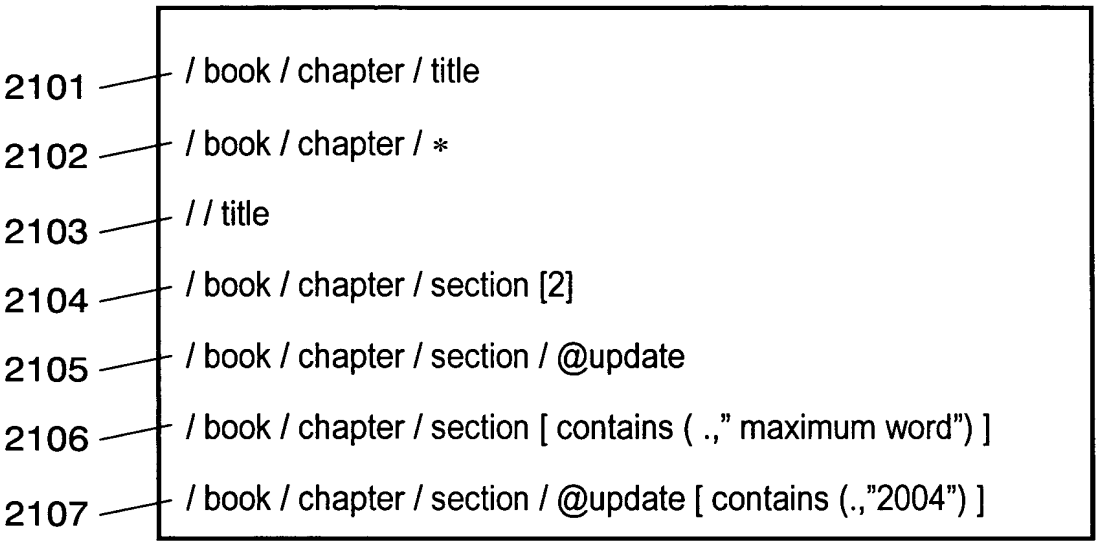
2101
/ book / chapter / title
2102
/ book / chapter / *
2103
/ / title
2104
/ book / chapter / section [2]
2105
/ book / chapter / section / @update
2106
/ book / chapter / section [contains (.," maximum word")]
2107
/ book / chapter / section / @update [contains (.,"2004")]

FIG. 15

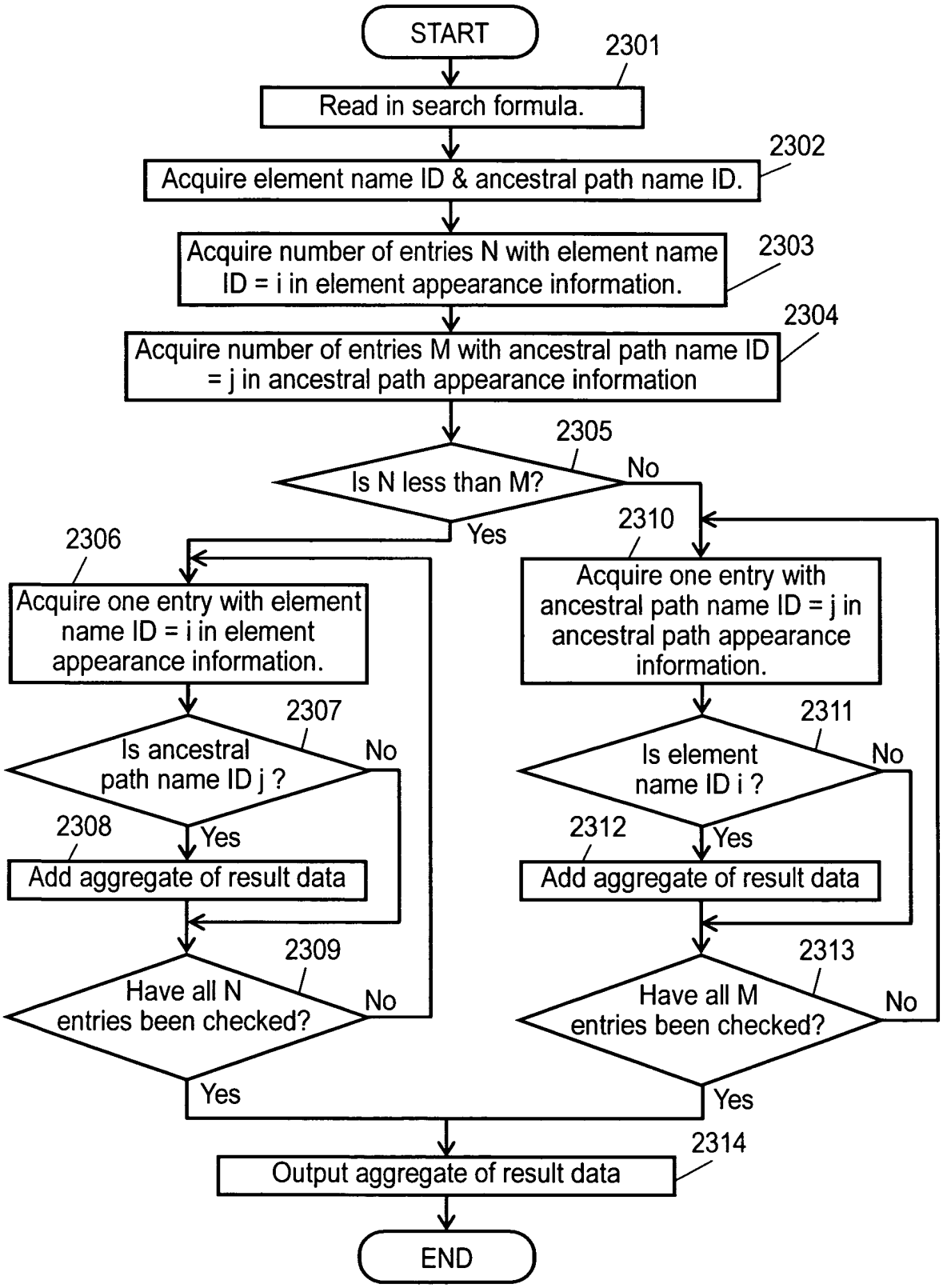
START
2301
Read in search formula.
2302
Acquire element name ID & ancestral path name ID.
2303
Acquire number of entries N with element name ID = i in element appearance information.
2304
Acquire number of entries M with ancestral path name ID = j in ancestral path appearance information
2305
Is N less than M?
No
Yes
2306
Acquire one entry with element name ID = i in element appearance information.
2307
Is ancestral path name ID j ?
No
Yes
2308
Add aggregate of result data
2309
Have all N entries been checked?
No
Yes
2310
Acquire one entry with ancestral path name ID = j in ancestral path appearance information.
2311
Is element name ID i ?
No
Yes
2312
Add aggregate of result data
2313
Have all M entries been checked?
No
Yes
2314
Output aggregate of result data
END

FIG. 16A

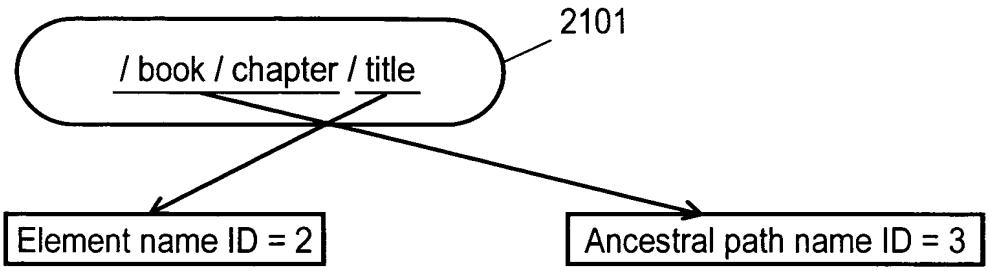

FIG. 16B

Element name ID	Document NO.	Character position	Number of characters	Ancestral path name ID	Order of branches
	⋮				
2	1	0	4	2	1/1
	1	4	7	3	1/1/1
	1	59	4	3	1/2/1
	2	0	6	2	1/1
	4	0	8	2	1/1
	4	24	6	3	1/1/1
	7	0	5	2	1/1
	9	7	4	3	1/1/1
3	⋮				

1301

FIG. 16C

(Document No., Ancestral path name ID, Element name ID, Attribute name ID, Order of branches)
={ (1 , 3 , 2 , 0 , 1/1/1),
(1 , 3 , 2 , 0 , 1/2/1),
(4 , 3 , 2 , 0 , 1/1/1),
(9 , 3 , 2 , 0 , 1/1/1) }

1302

FIG. 17A

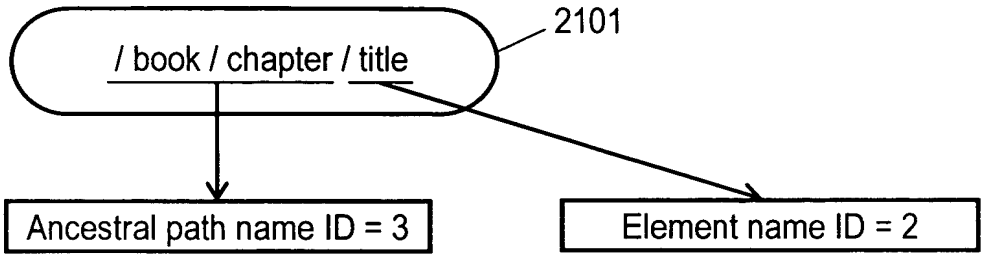

FIG. 17B

Ancestral path name ID	Document No.	Character position	Number of characters	Element name ID	Order of branches
	⋮				
3	1	4	7	2	1/1/1
	1	11	28	4	1/1/1
	1	39	20	4	1/1/2
	1	59	4	2	1/2/1
	1	63	30	4	1/2/1
	1	93	22	4	1/2/2
	1	115	40	4	1/2/3
	3	5	25	4	1/1/1
	4	24	6	2	1/1/1
	4	60	15	4	1/1/1
	6	64	6	4	1/1/1
	9	7	4	2	1/1/1
4	⋮				

1401

FIG. 17C

(Document No., Ancestral path name ID, Element name ID, Attribute name ID, Order of branches)
={ (1 , 3 , 2 , 0 , 1/1/1),
(1 , 3 , 2 , 0 , 1/2/1),
(4 , 3 , 2 , 0 , 1/1/1),
(9 , 3 , 2 , 0 , 1/1/1) }

1402

FIG. 18A

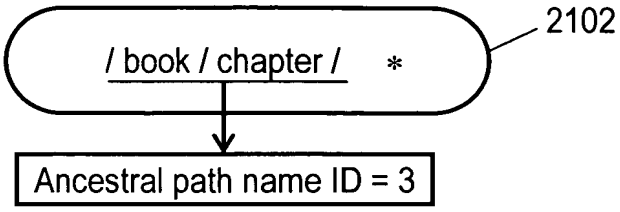

FIG. 18B

Ancestral path name ID	Document No.	Character position	Number of characters	Element name ID	Order of branches
	⋮				
3	1	4	7	2	1/1/1
	1	11	28	4	1/1/1
	1	39	20	4	1/1/2
	1	59	4	2	1/2/1
	1	63	30	4	1/2/1
	1	93	22	4	1/2/2
	1	115	40	4	1/2/3
	3	5	25	4	1/1/1
	4	24	6	2	1/1/1
	4	60	15	4	1/1/1
	6	64	6	4	1/1/1
	9	7	4	2	1/1/1
4	⋮				

1501

FIG. 18C

(Document No. Ancestral path name ID, Element name ID, Attribute name ID, Order of branches)
={ (1 , 3 , 2 , 0 , 1/1/1),
(1 , 3 , 4 , 0 , 1/1/1),
.....
(6 , 3 , 4 , 0 , 1/1/1),
(9 , 3 , 2 , 0 , 1/1/1) }

1502

FIG. 19A

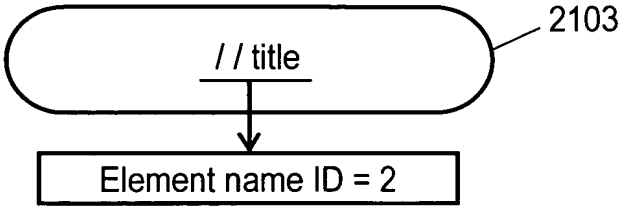

FIG. 19B

Element name ID	Document No.	Character position	Number of characters	Ancestral path name ID	Order of branches
	⋮				
2	1	0	4	2	1/1
	1	4	7	3	1/1/1
	1	59	4	3	1/2/1
	2	0	6	2	1/1
	4	0	8	2	1/1
	4	24	6	3	1/1/1
	7	0	5	2	1/1
	9	7	4	3	1/1/1
3	⋮				

1601

FIG. 19C

(Document No., Ancestral path name ID, Element name ID, Attribute name ID, Order of branches)
={ (1 , 2 , 2 , 0 , 1/1),
(1 , 3 , 2 , 0 , 1/1/1),
.....
(7 , 2 , 2 , 0 , 1/1),
(9 , 3 , 2 , 0 , 1/1/1) }

1602

FIG. 20A

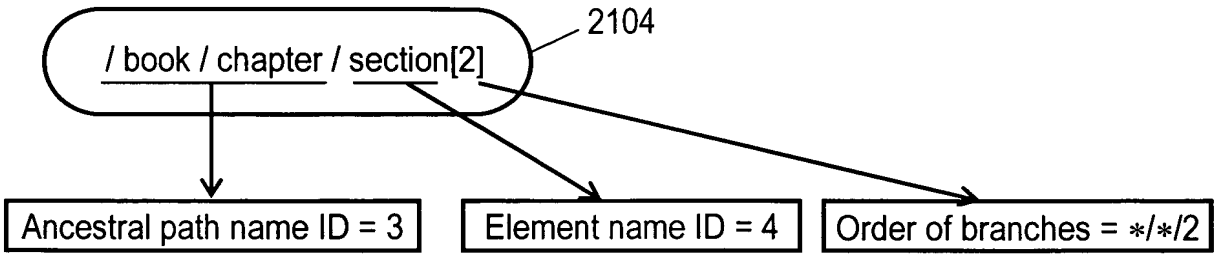

FIG. 20B

Ancestral path name ID	Document No.	Character position	Number of characters	Element name ID	Order of branches
		1	⋮		
3	1	4	7	2	1/1/1
	1	11	28	4	1/1/1
	1	39	20	4	1/1/2
	1	59	4	2	1/2/1
	1	63	30	4	1/2/1
	1	93	22	4	1/2/2
	1	115	40	4	1/2/3
	3	5	25	4	1/1/1
	4	24	6	2	1/1/1
	4	60	15	4	1/1/1
	6	64	6	4	1/1/1
	9	7	4	2	1/1/1
4			⋮		

1701

FIG. 20C

(Document No., Ancestral path name ID, Element name ID, Attribute name ID, Order of branches)
={ (1 , 3 , 4 , 0 , 1/1/2),
(1 , 3 , 4 , 0 , 1/2/2) }

1702

FIG. 21A

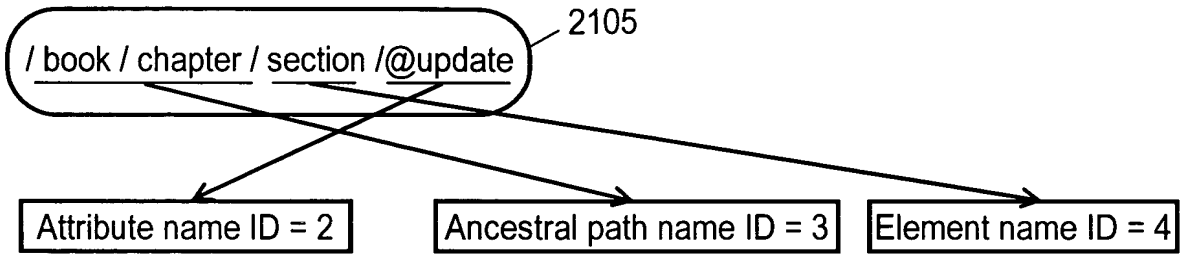

FIG. 21B

Attribute name ID	Document No.	Character position	Number of character	Ancestral path name ID	Element name ID	Order of branches	
	⋮						
2	1	115	6	3	4	1/2/3	1801
	2	8	4	2	2	1/1	
	5	60	6	3	4	1/1/2	
	8	32	8	3	2	1/2/1	
3	⋮						

FIG. 21C

(Document No., Ancestral path name ID, Element name ID, Attribute name ID, Order of branches)
={ (1 , 3 , 4 , 2 , 1/2/3),
(5 , 3 , 4 , 2 , 1/1/2) } 1802

FIG. 22A

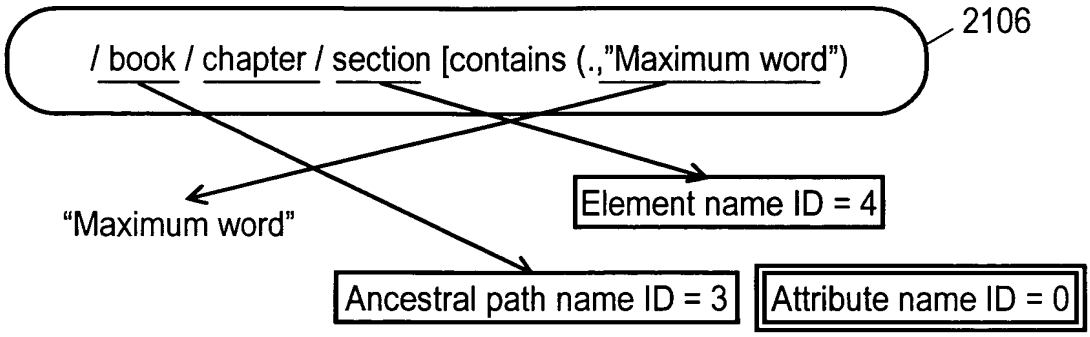

FIG. 22B

Partial character string	Document No.	Character position	Ancestral path name ID	Element name ID	Attribute name ID	Order of branches	
			⋮				
"Maximum"	1	118	3	4	0	1/2/3	1901
	2	86	3	4	0	1/1/1	
	3	24	2	2	0	1/1	
	4	62	3	4	0	1/1/1	
	8	77	3	4	2	1/1/1	
	1904		⋮				
"Word"	1	120	3	4	0	1/2/3	1902
	3	26	2	2	0	1/1	
	4	64	3	4	0	1/1/1	
	8	79	3	4	1	1/1/1	
	1905		⋮				

FIG. 22C

(Document No., Ancestral path name ID, Element name ID, Attribute name ID, Order of branches)
={ (1 , 3 , 4 , 0 , 1/2/3),
(4 , 3 , 4 , 0 , 1/1/1) } 1903

FIG. 23A

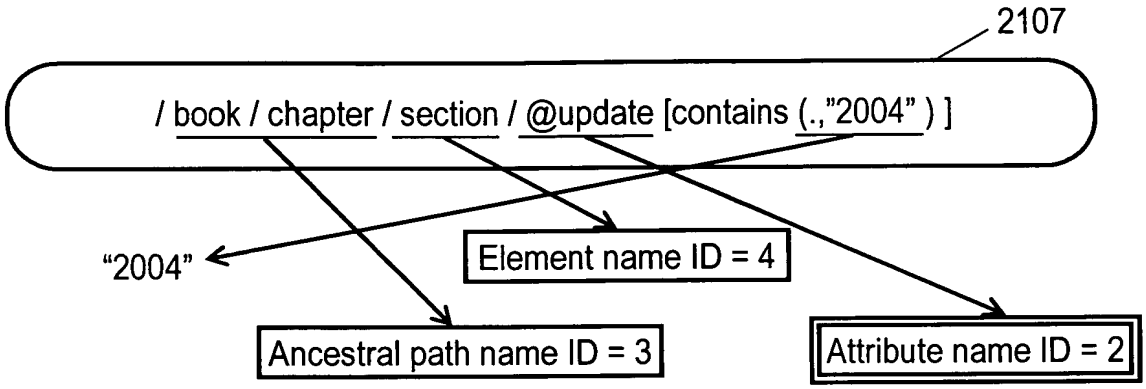

FIG. 23B

Partial character string	Document No.	Character position	Ancestral path name ID	Element name ID	Attribute name ID	Order of branches	
			⋮				
"20"	1	115	3	4	2	1/2/3	2001
	2	15	3	4	0	1/1/1	
	3	24	2	2	0	1/1	
	5	21	3	4	2	1/1/1	
	7	54	3	4	1	1/1/1	
	2004		⋮				
"04"	1	117	3	4	2	1/2/3	2002
	3	26	2	2	0	1/1	
	5	23	3	4	2	1/1/1	
	7	56	3	4	1	1/1/1	
	2005		⋮				

FIG. 23C

(Document No., Ancestral path name ID, Element name ID, Attribute name ID, Order of branches)
={ (1 , 3 , 4 , 2 , 1/2/3),
(5 , 3 , 4 , 2 , 1/1/1) } 2003

FIG. 24

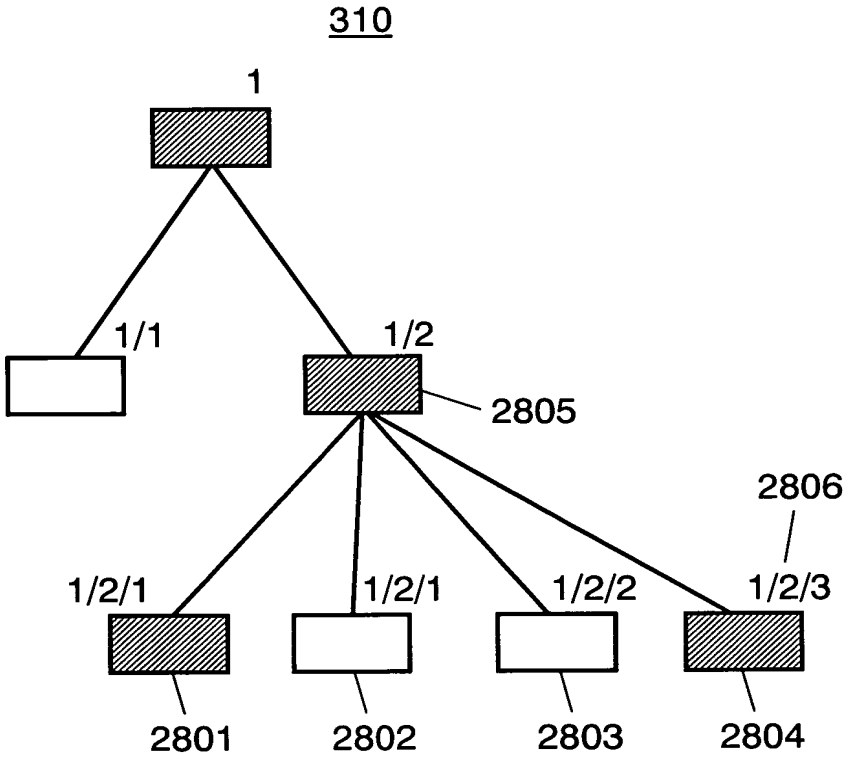

FIG. 25A

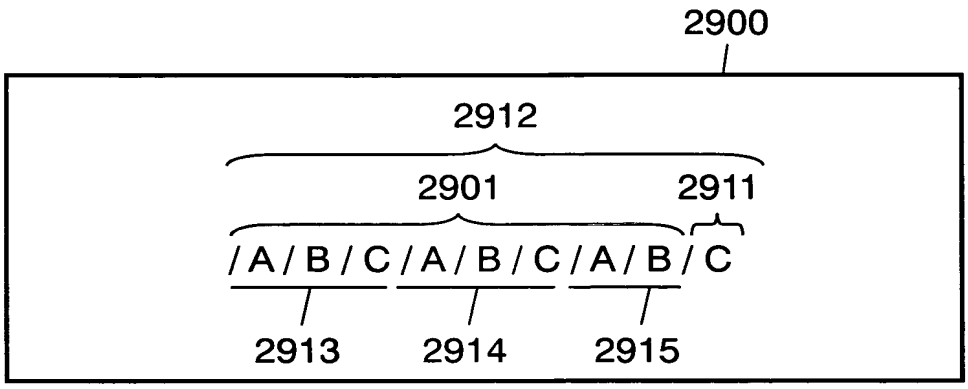

FIG. 25B

2903

2904 Ancestral path ID	2905 Ancestral path name
1	/
.....	.....
25	/A/B
.....	.....
83	/A/B/C
.....	.....

FIG. 25C

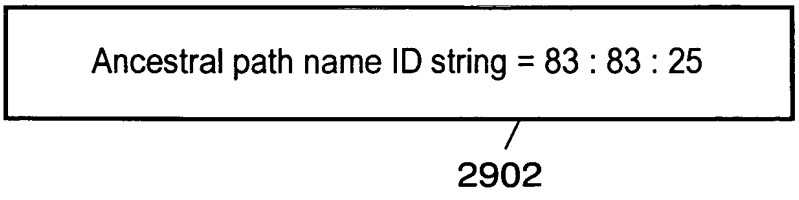

FIG. 26

541

502 Element name ID	503 Document No.	504 Character position	505 Number of characters	506 Ancestral path name ID	507 Order of branches	548 Order of empty elements
10	1	100	20	83 : 25	1/2/3/1/1/2	1/1/2/1/2/1

FIG. 27

551

506 Ancestral path name ID	503 Document No.	504 Character position	505 Number of characters	502 Element name ID	507 Order of branches	548 Order of empty elements
83 : 25	1	100	20	10	1/2/3/1/1/2	1/1/2/1/2/1

FIG. 28

/ A / B / X / following – sibling : : Y

3201

FIG. 29A

Ancestral path name ID	Document No.	Character position	Number of characters	Element name ID	Order of branches	Order of empty elements
			.....			
25			.....			
	2	80	10	10	1/1/1	1/2/1
			.....	Element name X		
	2	120	20	14	1/1/1	1/2/1
			.....	Element name Y		
	5	100	0	10	1/2/1	1/1/1
	5	100	30	14	1/2/1	1/2/2
			.....			
26			.....			

3301

3302

FIG. 29B

(Document No., Ancestral path name ID, Element name ID, Attribute name ID, Order of branches, Order of empty elements)
={ (2 , 25 , 14 , 0 , 1/1/1 , 1/2/1),
(5 , 25 , 14 , 0 , 1/2/1 , 1/1/2) }

3303

FIG. 30

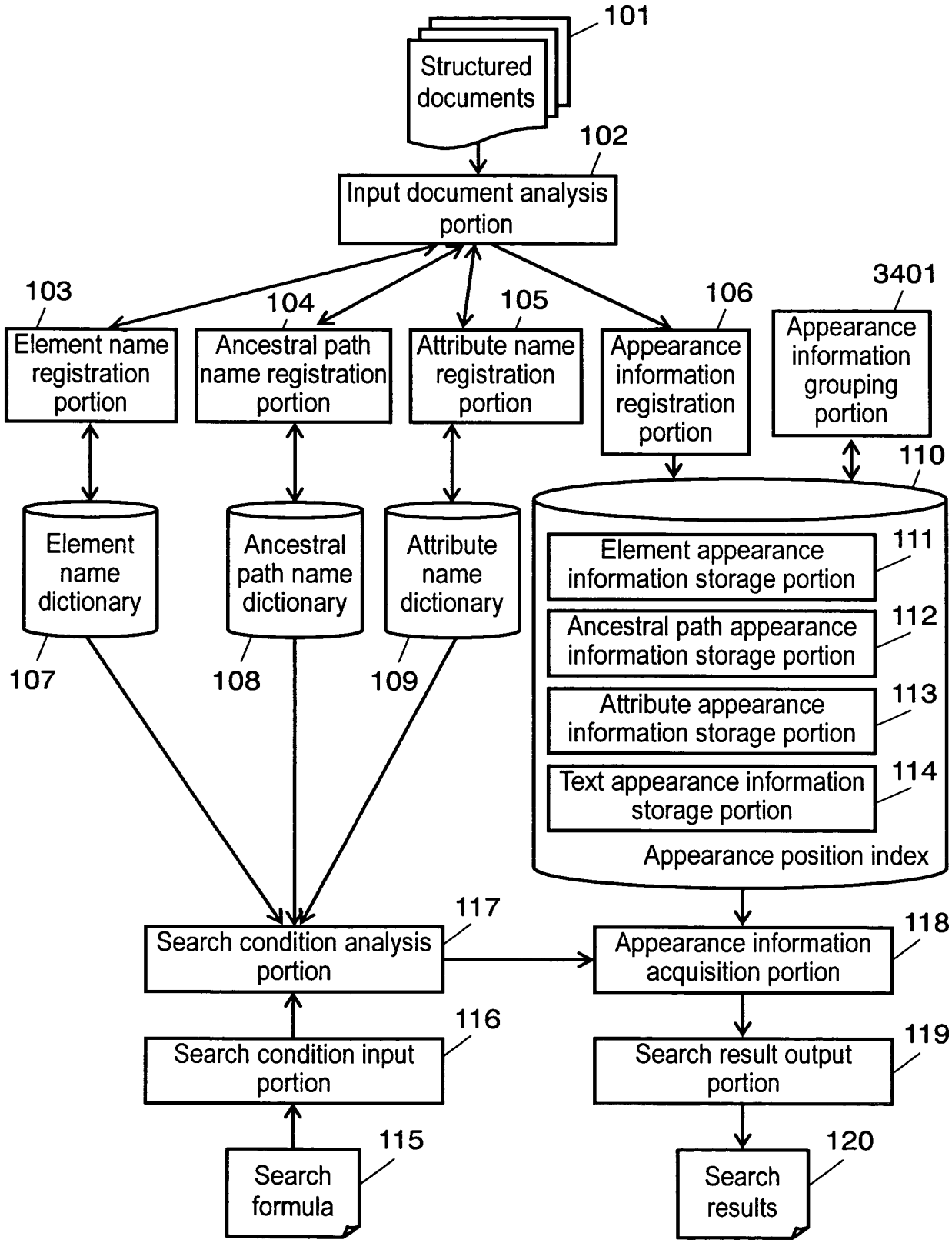
101
Structured documents
102
Input document analysis portion
103
Element name registration portion
104
Ancestral path name registration portion
105
Attribute name registration portion
106
Appearance information registration portion
3401
Appearance information grouping portion
110
Element name dictionary
Ancestral path name dictionary
Attribute name dictionary
107
108
109
Element appearance information storage portion
111
Ancestral path appearance information storage portion
112
Attribute appearance information storage portion
113
Text appearance information storage portion
114
Appearance position index
117
Search condition analysis portion
118
Appearance information acquisition portion
116
Search condition input portion
119
Search result output portion
115
Search formula
120
Search results

FIG. 31

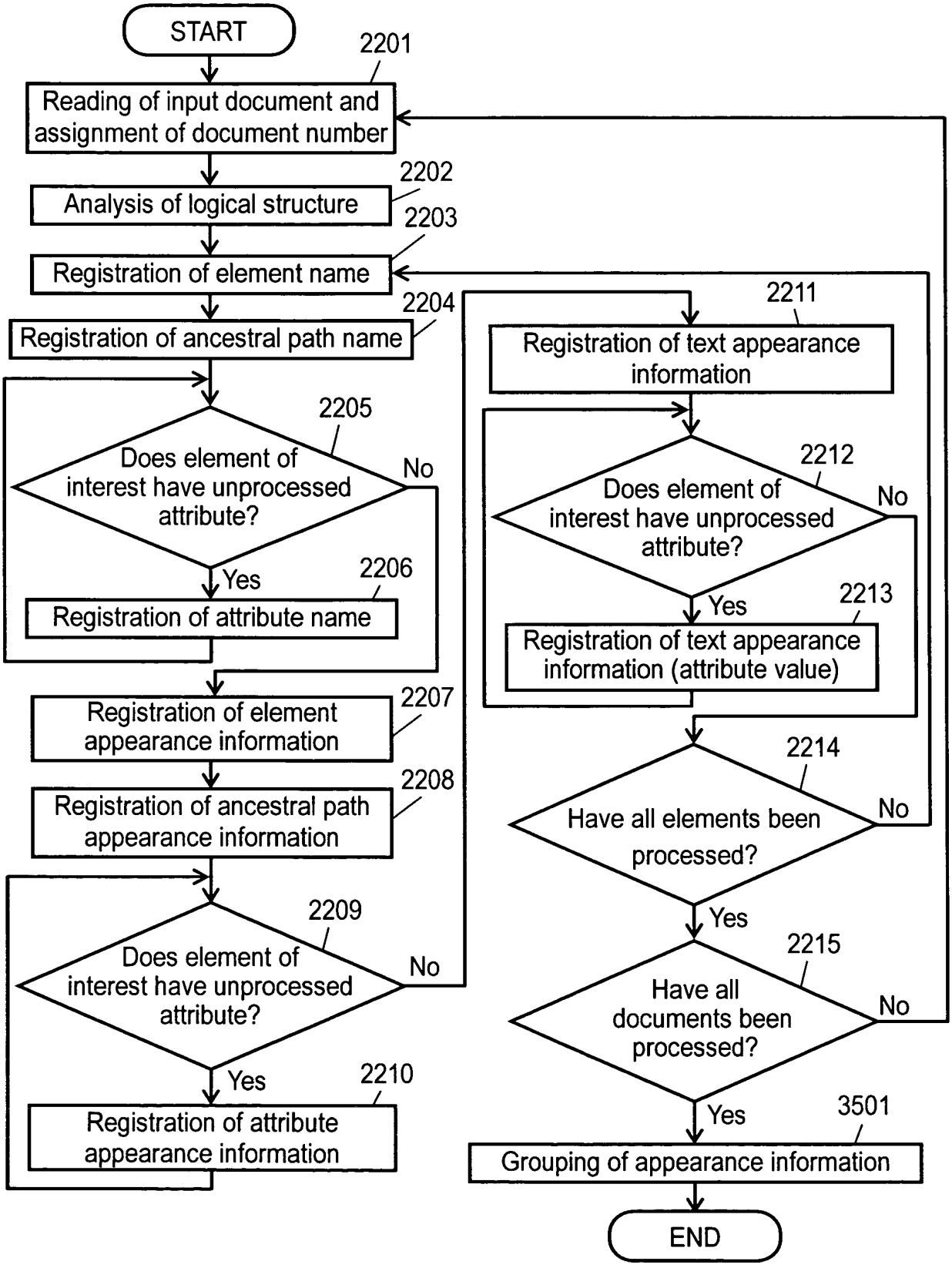
START
2201
Reading of input document and assignment of document number
2202
Analysis of logical structure
2203
Registration of element name
2204
Registration of ancestral path name
2205
Does element of interest have unprocessed attribute?
No
Yes
2206
Registration of attribute name
2207
Registration of element appearance information
2208
Registration of ancestral path appearance information
2209
Does element of interest have unprocessed attribute?
No
Yes
2210
Registration of attribute appearance information
2211
Registration of text appearance information
2212
Does element of interest have unprocessed attribute?
No
Yes
2213
Registration of text appearance information (attribute value)
2214
Have all elements been processed?
No
Yes
2215
Have all documents been processed?
No
Yes
3501
Grouping of appearance information
END

FIG. 32

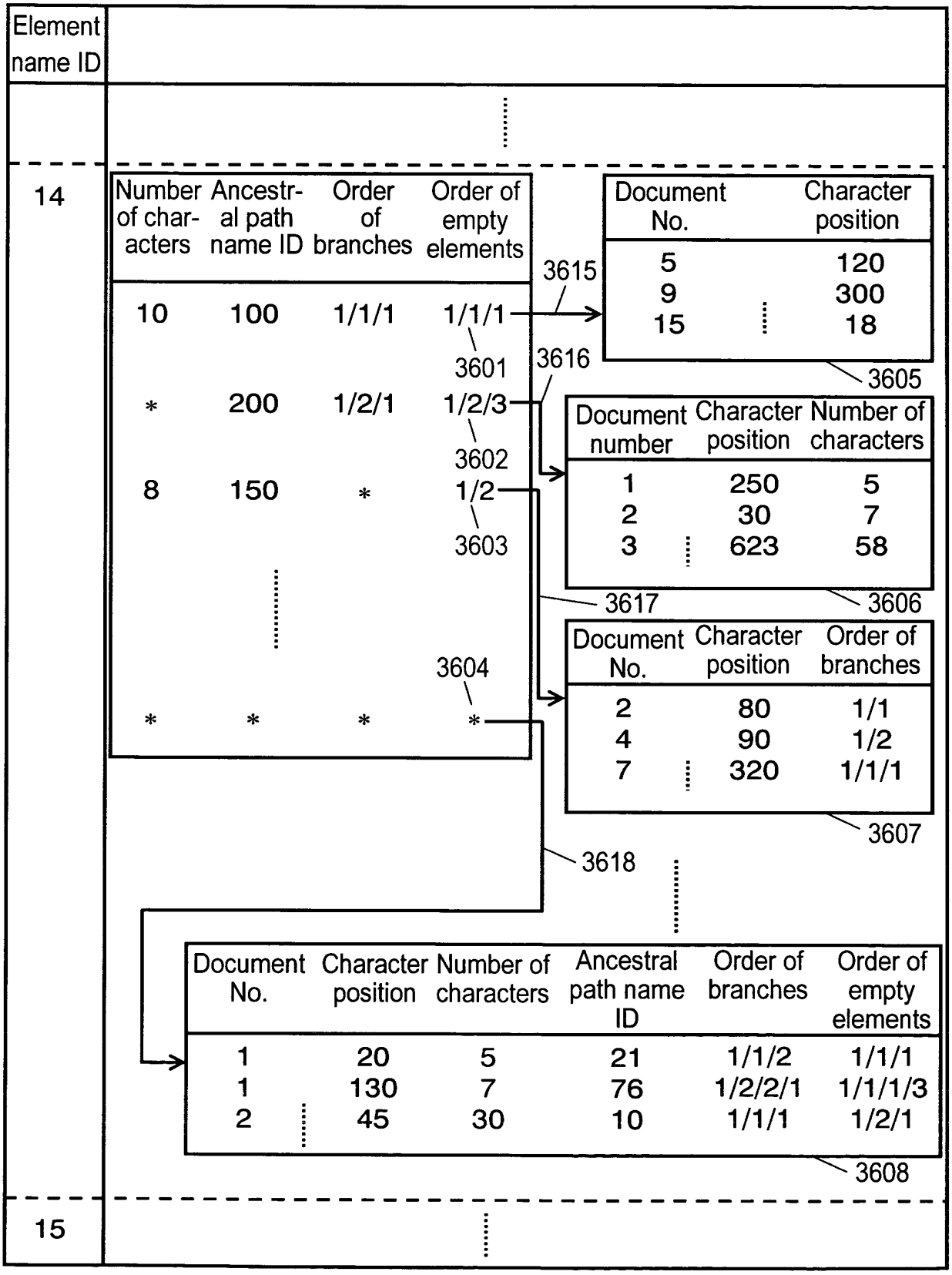
Element name ID
14
Number of characters
Ancestral path name ID
Order of branches
Order of empty elements
10
100
1/1/1
1/1/1
3601
*
200
1/2/1
1/2/3
3602
8
150
*
1/2
3603
3604
*
*
*
*
3615
Document No.
Character position
5
120
9
300
15
18
3605
3616
Document number
Character position
Number of characters
1
250
5
2
30
7
3
623
58
3606
3617
Document No.
Character position
Order of branches
2
80
1/1
4
90
1/2
7
320
1/1/1
3607
3618
Document No.
Character position
Number of characters
Ancestral path name ID
Order of branches
Order of empty elements
1
20
5
21
1/1/2
1/1/1
1
130
7
76
1/2/2/1
1/1/1/3
2
45
30
10
1/1/1
1/2/1
3608
15

FIG. 33

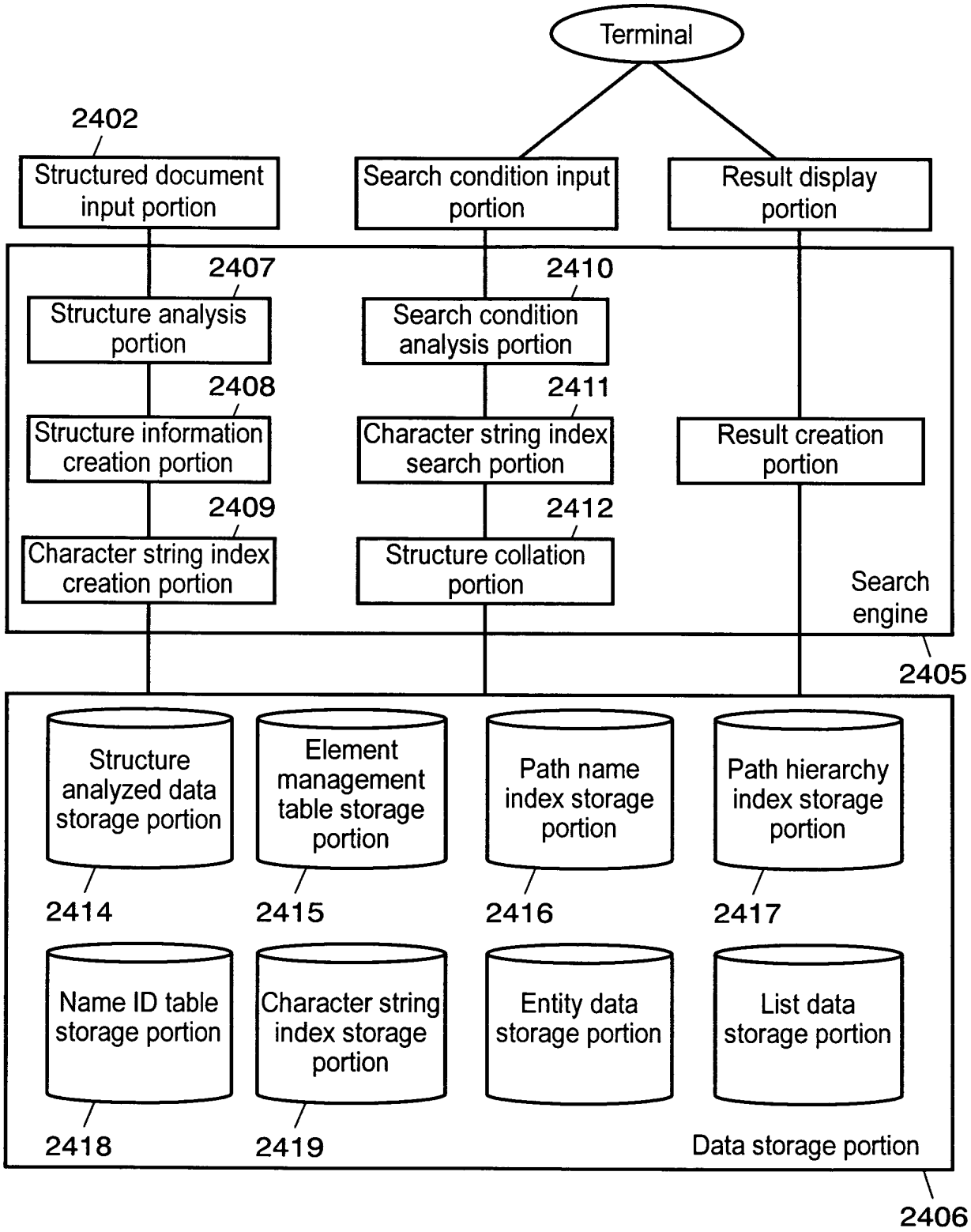
Terminal
2402
Structured document input portion
Search condition input portion
Result display portion
2407
Structure analysis portion
2410
Search condition analysis portion
2408
Structure information creation portion
2411
Character string index search portion
Result creation portion
2409
Character string index creation portion
2412
Structure collation portion
Search engine
2405
Structure analyzed data storage portion
Element management table storage portion
Path name index storage portion
Path hierarchy index storage portion
2414
2415
2416
2417
Name ID table storage portion
Character string index storage portion
Entity data storage portion
List data storage portion
2418
2419
Data storage portion
2406

FIG. 34

2501

2502 Search unit identifier	2503 Document No.	2504 Path name ID	2505 Path hierarchy ID	2506 Name ID
1	1	N2	L2	T3
2	1	N3	L2	T4
3	1	N3	L6	T4
4	1	N4	L2	T5
5	1	N7	L3	T8
6	1	N8	L3	T9
7	1	N10	L4	T11
8	1	N11	L4	T9
⋮	.....	.....	.....	.....

FIG. 35A

< title > structured document management < /title >

FIG. 35B

2602

2606	2603 Chain of characters	2604 Search unit identifier	2605 Character position No.
	structur-	1	1
	-red	1	2
	-ed do-	1	3
	document	1	4
	ment mana-	1	5
	management	1	6

FIG. 36A

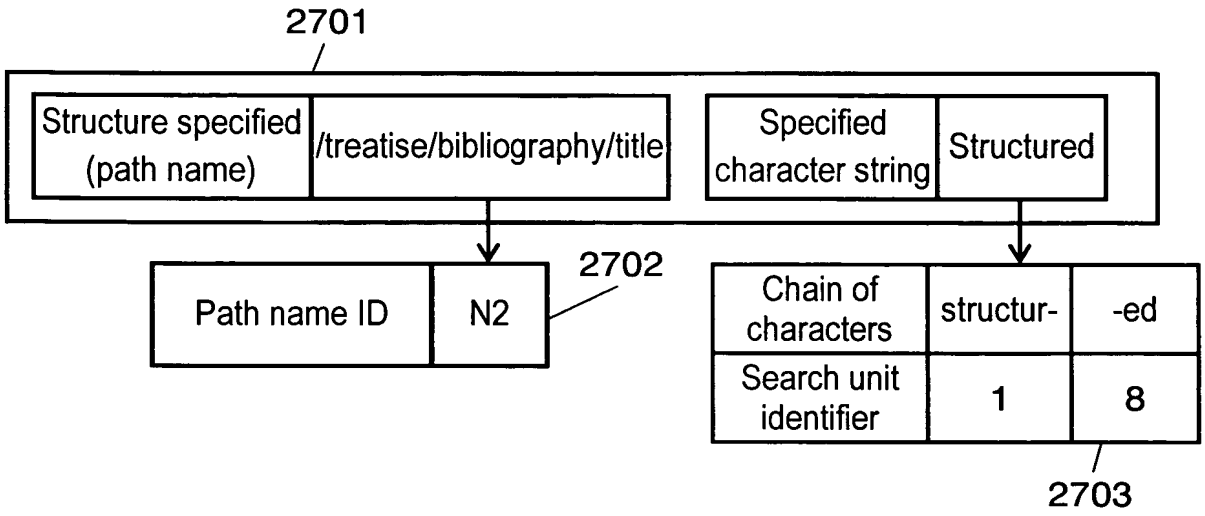

FIG. 36B

2501

Search unit identifier	Document No.	Path name ID	Path hierarchy ID	Name ID
1	1	N2	L2	T3
⋮	.....	.....	.....	.....
8	1	N11	L4	T9
⋮	.....	.....	.....	.....

FIG. 36C

Search unit identifier	Document No.
1	1